DOCUMENT RESUME

ED 193 321 TM 800 638

AUTHOR Hubert, John A.
TITLE Problems in Language Proficiency Assessment: A Review of the Language Assessment Battery (English, Level I).
PUB DATE Oct 80
NOTE 29p.: Paper presented at the Annual Meeting of the Northeastern Educational Research Association (October, 1980).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Audiolingual Skills; *Bilingual Education; Hispanic Americans; Item Analysis; *Language Proficiency; *Language Tests; Primary Education; Reading Achievement; Screening Tests; *Spanish Speaking; Test Reliability; *Test Validity; Writing (Composition)
IDENTIFIERS Connecticut; *Language Assessment Battery

ABSTRACT

Questions are addressed pertaining to the Language Assessment Battery (LAB), currently the only single battery that meets all of the state of Connecticut criteria for an instrument to determine which pupils must be offered bilingual education. Normed on an English dominant population, the primary purpose was to obtain its psychometric characteristics for pupils of limited English proficiency. Other purposes were to determine its selection accuracy, and to examine it as a measure of English achievement within a bilingual program. Validity was also reviewed. The psychometric portion of the study used 195 K-2 Hispanics in a bilingual program. Scale structure, internal consistency reliability, and item statistics were analyzed. The other investigations used the publisher's information. The problems encountered were instructive for the makers of this type of test. Psychometrically the test is sound except for the low reliabilities and poor item statistics of the reading and writing subscales in K and grade 1. Those scales at these grades also have little content validity. Standard error of measurement is very poor in terms of percentile rank near the cut scores used for selection. Establishment of standards of construct validity for this new type of test are recommended. (Author/GK)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED193321

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

PROBLEMS IN LANGUAGE PROFICIENCY ASSESSMENT: A REVIEW OF THE LANGUAGE ASSESSMENT BATTERY (ENGLISH, LEVEL I)

John A. Hubert

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J. Hubert

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Presented at the Annual Meeting of the
Northeastern Educational Research Association
October, 1980

TM 800 638

PROBLEMS IN LANGUAGE PROFICIENCY ASSESSMENT: A REVIEW OF THE LANGUAGE ASSESSMENT BATTERY (ENGLISH, LEVEL I)

Problem Statement

The New York City Board of Education undertook the development of the Language Assessment Battery (LAB) in 1974 for the express purpose of obtaining "an improved method of classifying ...children whose English language deficiencies prevent them from effectively participating in the learning process ... "i.e., to help decide which children should be enrolled in bilingual programs. The full battery so developed has English and Spanish versions on three levels spanning grades K-12. Subtests include Listening, Speaking, Reading, and Writing. Level I combines Listening and Speaking into one, making three in all. The Level I Battery is individually administered.

The English battery of the LAB has become particularly important because it is so far the only single battery accepted as valid by the Connecticut State Department of Education for the purpose of proficiency assessment in screening and placement procedures across grades K-12.

Hartford Public Schools began using the LAB as one of the steps in its screening and placement procedures in 1977-1978. It was also selected as a pre-post measure of English language achievement in a Title I and Title VII

program. This was done primarily to avoid two additional testings in a second individually administered instrument. During the course of its use in the fall and winter, a number of questions arose which focused most strongly on the English Level I. Those questions addressed here are:

1)What are the psychometric properties of the LAB when administered to students who are pre-selected for limited English proficiency?

2) What is the screening accuracy of the LAB in the vicinity of the cut-off score?

3) Can the LAB be used to assess achievement gains?

Test Content

The Language Assessment Battery is a new type of instrument. It was commissioned as a result of a consent decree (72 Civ 4002) on behalf of Hispanic school children. The decree noted "that both expressive and receptive skills need to be assessed and that there were no published instruments which adequately perform this task". (LAB Technical Manual)

Despite the fact that this is a relatively new type of instrument, the LAB is apparently designed considering content validity, i.e. "the extent to which each test constitutes a representative sample of the skills, knowledge and understanding of the goals of instruction..." It also generally follows the pattern of development used for achievement tests by following objectives and blueprints, etc.

The manual lists the objective for each item, a notable practice. However each objective seems to be scarcely more than a re-statement of the item. The list of objectives (Table I) is not particularly reassuring that the items constitute a reasonable sample of the skills constituting language proficiency. Six of twenty items deal with parts of the body, but weather, clothing, days of the week and many other topics are not covered. These items seem to have been chosen in part for the ease with which they could be administered. For instance, the examiner can easily ask the pupil to name the part of the body touched, whereas more representative questions may be (or were thought to have been) more ambiguous to convey.

On the basis of this general survey of content, one can easily argue that the failure to know these items tends to confirm an absence of English proficiency. But the difference between pre- and post-test scores is not necessarily a good assessment of a program of English instruction because of limited sampling of the usual content in these grades.

Test administration requires the administration manual, a Picture Stimulus Booklet for Listening & Speaking and for Reading, a Student's Test Booklet for Writing, and an answer sheet for each pupil completed by the examiner.

At the beginning of each section are verbal instructions. It is assumed that these instruction are to be given in English. These instructions raise some questions, however, because their vocabularly and grammar are more complex than the items. It is therefore unclear whether they are intended to be part of the test, and whether in fact they are part of the test. It is not clear, whether the general instructions as opposed to the item stimuli could be given in the child's first language.

The pictures in the stimulus booklet are large and clear. With exceptions as noted below, they generally seem appropriate for Kindergarteners in this respect.

In Listening & Speaking the first three questions are: What's your name? How old are you?, and What grade are you in? Any appropriate verbal answer in English is correct.

Items 4-17 deal with basic vocabulary (nouns). The child must touch a part of the body or selected pictures of colors or common objects in response to instruction; or the child must name the part of the body or picture of the object indicated by the examiner.

Items 18,19 are more complex in format. The child is told to "touch the picture which shows what I say". In 18, "the cat and three kittens are playing". The distractors are a boy and girl riding bicycles, and an airplane and bird flying. The recognition of a single

word will still clue the correct answer. The same is true in 19 where the child must select a green circle from a yellow square and red triangle.

Item 20 continues the format but each picture involves a girl and puppy. The child must know what the puppy is doing, namely, following her. The distractors place the puppy in her arm and between her feet. The last distractor is not so attractive since a young child taking the test might see the puppy between her feet as following her. A following puppy often winds up there. The adequacy of the distractor cannot be checked with standard scoring since only right or wrong is recorded for any item. A record is not kept of the distractor chosen.

The Reading subtest has ten items (21-30). Each item has a picture and either three words or three sentences. The child is instructed: "Here is a picture and three words (sentences). Which word (sentence) says what the picture shows?" Pictures Include an apple, boy climbing a fence, elephant eating hay, sleeping puppy, street with apartment houses, two bears playing a grand piano, girl and dog in the park. Each of the first three pictures are used twice.

Generally the stimuli, responses, and distractors seem appropriate as items for beginning readers, but not

for Kindergarteners and early first graders. However, repeated use of the same stimulus does not seem right when there are so few items, and when the test is not measuring materials in which the pupil has recently received instruction. It is poor sampling of the possible range of materials.

This objection seems particularly strong against the elephant because some children may not be familiar with this animal and thereby miss two items. Also, the second use of the item for 'eating' seems relatively obscure for this word, in light of the general level of the test and its purpose.

As in the Reading subtest, the section on Writing is inappropriate for Kindergarten and early first grade. Failure is no indication of lack of fluency in English. It is more likely to be an index of social class than language fluency.

The first two items each ask the pupil to copy a letter of the alphabet. This skill is certainly a writing skill, but it is not specific to any particular language. Perhaps these items were intended as an easy start. In this author's opinion, failure on either item is adequate reason to discontinue the section, but no directions to that effect are given.

Items 33-35 call for the child to complete a word so as to describe a given picture: The child is given letters to choose from:

(boat)	-oat	b c r v
(party)	part-	h d y e
(school)	sch--l	o u oo aa

It seems to me possible that fluent first graders and slow second graders could be confused by the assignment. Why not simply ask children to write the word "boat" or write a word to go with the picture? On the positive side, however, it appears that the distractors might have been designed with the Hispanic in mind, who might be tempted to spell voat, parte, or schul because of the characteristics of the Spanish language. These items are interesting because they suggest the possible importance of viewing English skills from the point of view of the non-English learner. It also raises the interesting idea of an English test whose validity is limited to specific other- language learners, in this case Spanish-speaking.

Items 36-38 give a picture and four words. The child is directed to write a word to go with the picture, one of the four. The word to be written is also spoken by the examiner. The only problem with these items is that they seem to test reading more than writing. Being able to copy is not the same as writing.

Items 39-40 are unusual. A picture is given (picture of a bird's nest) and the following sequence of words:

eggs nest are the in the

The child is told: "The words are out of order; put them in order so they tell what is in the picture".

Again some fluent children will know the right word order, but other fluent children may be confused by the assignment. The item seems to test logical skills more than writing proficiency.

Item 40 has some troubling features. The grocer is weighing food which appears to be grapes. The backdrop further establishes the setting as the produce department at the market. A girl is watching; it is not clear that she is holding paper money and a child may not see it. The pupil may get the impression of the grocer weighing, not the girl buying.

With these two items, a question may be raised, namely that word order is idiosyncratic to each language, and that novel item types may be relevant. In this instance, though, the Spanish word order would be the same.

Psychometric Properties

The LAB was normed and validated on a sample of schools 75% of whose students were fluent in English. The important question to raise is whether the reliabilities and other statistics will remain acceptable when the test sample has a severely restricted range, i.e. the limited-English proficient students the test is designed for.

Item scores were extracted from 195 pre-tests for children in grades K-2. Internal reliability coefficients (alphas) were obtained for each scale and total score, for each grade and for total group. (Table I)

Correlations between each item and other items in the subscale were obtained for each grade and for total group, for each of the three subscales (Table II). Item means and standard deviations were also computed for each grade (Table III)

Principal component analyses were conducted using and oblique rotation. Each scale was analyzed separately,

and used all grade levels at once. Other subsets of the data were also factor analyzed, and described, but no data are reported.

The published norm tables were graphed (Figure 1), and other basic statistics were listed (Table VII). Then, error bands were calculated with respect to 23%ile, the score used for a cut-off in Hartford (Table VIII). Figure 1 and Table VIII address the feasibility of using this instrument for pupil placement decisions.

Results

Internal reliability coefficient alphas are reported in Table I below. Total score alphas for each

TABLE I

Internal Reliability Coefficient Alpha

Grade	N	Spk & List	Reading	Writing	Total
K	69	.910	.757	.518	.909
1	71	.838	.765	.542	.869
2	55	.859	.856	.742	.908
Total	195	.925	.869	.809	

grade are high, as are total group alphas for Speaking & Listening and for Reading. The total group reliability for writing is marginal at .809.

Alphas for individual grade and subscale vary widely. Four are quite high considering the number of items. Three are near .750 which is somewhat low relative to the standard of .80 but sturdy considering the restriction of the sample, both by English fluency and grade level. All of the foregoing alphas are similar to those found in the Technical Manual for the unrestricted norm sample. However, two alphas, for K and 1 in the Writing subtest, are low absolutely, and relative to the published data.

Item-total correlations for the total group are quite high for Speaking & Listening. Excepting the "give-away" item #1, they range from .49 - .76. Reading correlations range from .47 - .70 which is slightly lower but still quite high. Writing correlations vary from .37 - .58 for the total group, an acceptable standard of item-total correlations. See Table II.

Examined by grade, the correlations are lower, of course. A number of items have zero variance and produce zero correlations. For Kindergarten, the Speaking and Listening scale remains excellent. In Reading, six items are not answered and a seventh has a low correlation, leaving only three sound items. In writing, only two items appear sound (copying letters of the alphabet).

For the first grade there are several weak items in the first two scales, but no serious problems. The Writing

subscale, however, has but one item-total correlation above .4.

For the second grade there are two "dead" items in the first scale, none in the second, and one in the third. This last item (33) has virtually no item-total correlation at any single grade level, but is .5 over all.

Item means are reported in Table III. Scanning individual means shows which of the defective items (on an individual grade basis) are low scoring and which are high scoring. For instance in Grade 1, Writing (Table III), two items were answered correctly by all pupils, one by none, and several others by very few pupils.

Comparing item means across grades shows that in the Speaking and Listening subtest all but four items(5,6,8,12) are mastered early in the first grade. The second grade scores show no progress, and average total score has actually declined slightly. Progress has not been made on the four items, either. Thus Speaking and Listening is primarily relevant to discrimination between Kindergarten and first grade where most progress is recorded. Item means and total means for the scale indicate that the first grade is not teaching parts of the body and certain common words. (Since these are pre-test scores, the second grade scores report on first grade work.)

While it is possible that the failure of second grade Listening and Speaking scores to rise is real, it seems unlikely. This result is more probably attributable to the poor item sampling mentioned earlier. The absence of a gain is the result of grade-by-content sampling feature of the test which gives rise to a hidden ceiling effect. Sampling over a wider domain would have probably allowed scores to rise in a more conventional manner.

Growth across grades in the Reading test is uniform, about two points per grade average. In Writing most of the growth occurred between first and second grade. The scores are generally low because English reading and writing are not introduced until the second grade (after the pretest).

The component analysis revealed seven components that accounted for 59.2% of the variance. One component (#6) consisted of items 1, 2, (What is your name? and How old are you?). Each of the other subscales had two components. On examination of the item means in each component suggests that the two related components were separated primarily by difficulty level. The two components within Speaking & Listening were correlated at .51. Component correlation's were .28 and .15 for Reading and Writing respectively.

Several other subsets of the data were also analyzed to test the stability of the results. Analysis of the subscales separately, and of Reading and Writing together, using whole group and each grade separately, produced results very similar to that described above.

Checking all 40 items by grade produced spurious results because of the low pupil-to-variable ratio, which was less than 2:1.

Several items seem illogically placed when viewing content, and item 37 was unplaced. These discrepancies all occur on items with low loadings, and are probably idiosyncratic to the author's sample. The only exception is that items 31 and 32, copying letters, were at the head of the 'Oral language, easy' component. These two items are so easy compared with other writing tasks that they fell in with easy oral tasks (Component 3).

The statistical investigation confirms to a considerable extent the observations made in the review of the content:

-there is no point in including Reading and Writing subscales for Kindergarten and early first graders.

-The content sampling of the Listening and Speaking section is unsatisfactory.

In other repects the text has shown satisfactory reliability and item statistics. Further comments on scale structure follow.

Screening Characteristics

In NYC the cut score for selecting students for bilingual education was 20th %ile. In Hartford both the 23rd and 30th %ile were in use. For this particular investigation the 23rd %ile was used. For this analysis it matters little what cut score in this range is chosen.

Figure 1 is a graph of the norm tables in the Technical Manual. The Kindergarten curve is distributed normally. First grade is skewed negatively, and second grade more so. A curve much like the second graph was sought in the design so that most of the items would be effective in ranking accurately the language proficiency of pupils in the bottom 50%. That the test is utterly incapable of comparing superior second grade students is irrelevant.

Examination of the most favorable of the curves, second grade,shows that the aims were not entirely achieved. The test has the first requirement, namely, the median score is above 95% correct. However, 32 of 40, or 80% correct, is only at the 10%ile mark. Thus very few items separate 10%ile from 50%ile. Indeed a score of 25 ranks only 2%ile. The test is so easy that even a pupil below 1% can get most right.

The disadvantage of the steep normal curve is better seen in Table VII. The confidence intevals were calculated

using the publisher's data in Table VI. Basically there are few items in the 20-30%ile range, making it difficult to accurately rank scores in that vicinity.

Better discrimination in the bottom half could be attained by replacing a number of the easy items (e.g. 4, 10, 16, 18, 19, 30, 31, and 32) with items more difficult, but not so difficult as to lower the median. Items are needed that upper-half pupils can do consistently, but are more challenging than the current ones for the lower-half students.

Such a change would help the lower end of the Kindergarten and first grade curves as well. These too, waste too many items below the 10% level, and below the 1%ile level. The aim is a curve something like the one shown in Figure 2. Getting about 20% correct puts the pupil at the edge of the first percentile. The median is close to 100% correct, so that nearly 75% of the items are being used to discriminate across the bottom half of the scale.

Discussion and Conclusions

Level I of the English LAB when administered to students preselected for low English yielded favorable psychometric properties with important exceptions. Internal consistency reliability and item-rest correlations were very acceptable for the Listening & Speaking subscale at each grade level as was the reading subscale for grades 1 and 2. Reliability and item correlation data confirm the obvious content analysis that reading and writing are inappropriate content in kindergarten and that the writing subscale is providing no useful information in early first grade. The reading subscale produced satisfactory statisitcs in grades 1 and 2 despite the fact that few if any of the pupils had had English reading instruction.

It is apparent that either there should be a separate level for K and 1, or that the Reading and Writing subscales should not be administered at these grades. The intent of the consent decree would not be violated by this arrangement.

With the exception of the inappropriateness of some of the content at some grade levels, the LAB is psychometrically satisfactory for pupils preselected for low English skills.

The publisher's data have shown the test to have poor accuracy for screening in the vicinity of the usual cut-off

scores. This was determined by viewing the standard error of measurement in percentile terms. The cause of this problem was revealed by the component analyses and p-values in conjunction with the publisher's norm data. Careful attention to the clustering of p-values is needed in order to concentrate the discriminating power of the test near the cut-off score used for placement decisions.

At Level I, the English LAB should not be used to study achievement gains. The content sampled is limited with respect to topics taught; and the interplay of the restricted item sample with the curriculum used with the author's sample produced a hidden ceiling effect which could easily mar program evaluations. This may not be the case for other levels of the test which have many more items.

There were indications in the technical manual that the LAB was constructed at least in part on an achievement test model; for instance, there is a list of objectives and blueprint information. There are, however, some reasons to view this type of test with respect to construct validity.

One reason is the inclusion of unusual items. On the one hand, there were items not usually taught or tested, and which might be dependent on non-relevant issues such as I.Q.

More importantly, there were items which may(or may not) have beendesigned with respect to the learning characteristics of Hispanics acquiring English. This raises the question of whether items should be so designed. Such a test might include items to test the tendencies or errors a particular population would likely make with English, and careful attention to distractors would be needed. On the other hand making a test for all groups might be more marketable, though unequally difficult for different populations.

A second construct validity consideration is the relationship among the subscales. It has been argued that oral skills are the most important aspect of language proficiency (Hernandez-Chavez et.al.,1978) but that argument is made with respect to the construct in general. This test is used as a predictor of academic success on the assumption that the four constituent skills(listening, speaking, reading, writing) are equally important. But this is a testable hypothesis, and no evidence on the matter is given. This seems like a worthy topic for research.

TABLE II

CORRECTED ITEM-TOTAL CORRELATIONS

SPEAKING & LISTENING

	K	1	2	Total Group
1	.26	.25	-	.26
2	.57	.38	.58	.56
3	.52	.63	.39	.64
4	.66	.44	.66	.70
5	.36	.49	.51	.50
6	.54	.18	.39	.49
7	.68	.56	.33	.74
8	.60	.37	.45	.54
9	.66	.48	.18	.64
10	.60	.33	.54	.66
11	.37	.31	.58	.54
12	.42	.64	.50	.58
13	.78	.54	.59	.73
14	.57	.58	.46	.66
15	.49	.43	.65	.59
16	.59	-	.61	.62
17	.73	.65	.65	.76
18	.54	.52	.18	.55
19	.54	.17	-	.58
20	.48	.49	.61	.58

READING SUBSCALE

	K	1	2	Total Group
21	.77	.31	.58	.59
22	-	.44	.61	.62
23	.66	.47	.39	.52
24	-	.37	.40	.47
25	.66	.64	.56	.65
26	-	.34	.54	.52
27	-	.62	.61	.70
28	-	.46	.69	.66
29	.29	.42	.68	.62
30	-	.32	.55	.54

WRITING SUBSCALE

	K	1	2	Total Group
31	.56	-	.33	.40
32	.60	-	.33	.44
33	.01	-	.02	.50
34	.16	.03	.28	.37
35	-	.11	.49	.49
36	-	.28	.64	.58
37	.13	.33	.41	.53
38	-	.44	.43	.58
39	-	.29	.65	.56
40	-	-	.58	.51

TABLE III

ITEM MEANS & STANDARD DEVIATIONS

SPEAKING & LISTENING

	K		1		2	
Item	$\bar{X}$	S.D.	$\bar{X}$	S.D.	$\bar{X}$	S.D.
1	.96	.21	.99	.12	1.00	–
2	.71	.46	.88	.32	.96	.19
3	.41	.49	.80	.40	.89	.31
4	.62	.49	.99	.12	.95	.23
5	.16	.37	.48	.50	.51	.50
6	.17	.38	.46	.50	.69	.47
7	.35	.48	.90	.30	.93	.26
8	.19	.39	.61	.49	.45	.50
9	.55	.50	.90	.30	.91	.29
10	.55	.50	.96	.20	.93	.26
11	.55	.50	.90	.30	.96	.19
12	.17	.38	.59	.50	.55	.50
13	.36	.48	.75	.44	.84	.37
14	.39	.49	.83	.38	.84	.37
15	.58	.50	.90	.30	.87	.34
16	.68	.47	1.00	–	.93	.26
17	.49	.50	.89	.32	.89	.31
18	.77	.43	.97	.17	.96	.19
19	.67	.47	.99	.12	1.00	–
20	.65	.48	.93	.26	.91	.29
Total	9.99	5.53	15.72	3.29	14.96	3.33

READING

	K		1		2	
Item	$\bar{X}$	S.D.	$\bar{X}$	S.D.	$\bar{X}$	S.D.
21	.04	.21	.30	.46	.62	.49
22	–	–	.30	.46	.40	.49
23	.01	.12	.21	.41	.40	.49
24	–	–	.23	.42	.27	.45
25	.01	.12	.44	.50	.44	.50
26	–	–	.15	.36	.25	.44
27	–	–	.35	.48	.53	.50
28	–	–	.08	.28	.34	.48
29	.01	.12	.04	.20	.29	.46
30	–	–	.04	.20	.24	.43
Total	.09	.45	2.14	2.22	3.78	3.13

WRITING

	K		1		2	
Item	$\bar{X}$	S.D.	$\bar{X}$	S.D.	$\bar{X}$	S.D.
31	.65	.48	1.00	–	.95	.23
32	.55	.50	1.00	–	.95	.23
33	.04	.21	.55	.50	.78	.42
34	.04	.21	.17	.38	.44	.50
35	–	–	.10	.30	.36	.48
36	–	–	.18	.39	.40	.49
37	.03	.17	.35	.48	.55	.50
38	–	–	.28	.45	.53	.50
39	–	–	.03	.17	.25	.44
40	–	–	–	–	.25	.44
Total	1.32	1.01	2.66	1.43	5.45	2.39

3.66

TABLE IV

SUMMARY OF PRINCIPAL COMPONENT ANALYSIS
N = 195, Grades K-2

Item #	Item Task	Loading	Mean
	Component 1 - Oral Language, Difficult		
12	Name the object in the picture.	.89	.43
13	Name the object in the picture.	.65	.64
5	Touch the part of your body I name.	.64	.37
3	What grade are you in?	.62	.69
14	Name the object in the picture.	.60	.68
17	Touch the thing I name.	.54	.75
8	Name the parts of the body I touch.	.53	.42
33	Selects letter to complete a word.	.51	.44
16	Touch the thing I name.	.39	.87
	Component 2 - Writing, Difficult		
39	Makes sentence from scrambled words.	.88	.08
40	Makes sentence from scrambled words.	.85	.07
26	Selects word to match picture.	.56	.13
36	Writes word to match picture.	.49	.18
35	Writes letter to complete a word.	.39	.14
	Component 3 - Oral Language, Easy		
31	Reproduce the letter B.	.92	.86
32	Reproduce the letter E.	.83	.83
19	Touch the picture I name.	.75	.88
10	Touch the picture I name.	.72	.81
4	Touch the part of your body I name.	.62	.85
20	Touch the picture I name.	.48	.83
7	Name the part of the body I touch.	.48	.71
11	Touch the picture I name.	.47	.79
15	Touch the picture I name.	.44	.78
18	Touch the picture I name.	.42	.90
9	Name the part of the body I touch.	.37	.78
	Component 4 - Reading, Easy		
23	Select the word that matches the picture.	.78	.19
25	Select the word that matches the picture.	.66	.29
24	Select the word that matches the picture.	.61	.16
21	Select the word that matches the picture.	.60	.30
27	Select the word that matches the picture.	.54	.28
38	Writes word to match picture.	.37	.25
	Component 5 - Writing, Easy		
34	Writes letter to complete a word.	.70	.20
	Component 6 - Basic		
1	What is your name?	.76	.98
2	How old are you?	.58	.85
	Component 7 - Reading, Difficult		
28	Match sentence with picture.	.61	.13
30	Match sentence with picture.	.60	.08
29	Match sentence with picture.	.52	.10
6	Name the part of the body I touch.	.51	.43
22	Selects the word to match picture.	.44	.22

TABLE V
FACTOR PATTERN MATRIX FOR GRADES K-2

	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	FACTOR 6	FACTOR 7
1	-0.08357	-0.00578	0.07116	-0.7894	-0.07059	0.76403	0.03036
2	0.31832	-0.01522	0.11193	-0.02881	0.07931	0.58031	0.01213
3	0.61839	0.00297	-0.00152	-0.02093	0.21272	0.28197	0.01761
4	0.18986	0.05417	0.62487	0.04334	0.00959	0.19376	-0.04620
5	0.63895	0.19247	-0.06073	0.01993	-0.12119	-0.02307	0.13795
6	0.35511	-0.21660	0.25611	-0.02438	-0.00959	-0.12188	0.50892
7	0.37991	-0.09677	0.48498	-0.08770	0.22796	0.04347	0.04043
8	0.53148	0.07593	0.18172	0.06564	-0.03861	-0.16484	0.18710
9	0.28744	-0.00747	0.36934	-0.08475	0.08319	0.24635	-0.05134
10	0.07875	-0.00354	0.72419	-0.10659	0.05200	-0.04547	-0.01264
11	0.27817	-0.02043	0.47139	-0.11818	0.16076	-0.28793	-0.07562
12	0.69304	0.01765	-0.05663	-0.18924	-0.13682	-0.07497	0.15640
13	0.64629	-0.03028	0.11087	-0.13584	0.15421	0.14162	-0.04840
14	0.60304	0.01714	0.06481	-0.12210	0.13214	0.16582	-0.07951
15	0.37679	0.15575	0.43695	-0.02380	-0.21354	-0.11559	-0.13869
16	0.38524	0.06834	0.35526	0.12058	-0.00900	0.19142	-0.02222
17	0.54307	0.02024	0.29816	0.00706	0.06610	0.21669	-0.05366
18	0.20446	0.14257	0.42381	0.08831	-0.25336	0.32907	-0.07854
19	-0.06828	0.02876	0.75159	-0.06736	0.01668	0.09304	-0.02833
20	0.18579	0.07961	0.47758	-0.09290	-0.19704	0.14523	-0.08288
21	-0.02863	0.33702	0.09341	-0.59610	0.21838	-0.08919	-0.21300
22	0.01311	0.10028	0.08843	-0.38542	-0.05734	0.03495	0.43672
23	0.01098	0.00900	0.03687	-0.77596	-0.02875	0.01109	-0.05224
24	-0.05196	-0.10931	-0.00572	-0.60514	0.17921	0.15105	0.15815
25	0.06680	0.06765	0.05116	-0.66261	-0.20445	0.02450	0.20283
26	0.03335	0.56236	0.10141	-0.20465	-0.22886	-0.17150	0.08104
27	0.27383	0.12029	-0.02970	-0.54356	-0.05303	-0.02928	0.20240
28	0.02428	0.26332	-0.04077	-0.15613	0.19311	0.06843	0.60696
29	-0.10033	0.27096	0.02412	-0.27782	-0.05410	0.06763	0.51089
30	0.01018	0.32405	0.00928	0.02910	0.11717	0.03027	0.59571
31	-0.21889	-0.01826	0.91574	0.00399	0.05206	0.03032	0.12112
32	-0.09884	-0.02077	0.82910	0.01146	0.06549	0.02743	0.14193
33	0.51039	-0.01720	0.03680	-0.40974	0.04973	-0.03999	-0.04734
34	-0.00792	0.06063	0.16422	-0.02431	0.69598	-0.04847	0.06522
35	0.12667	0.39498	-0.04930	0.05489	0.35811	0.04637	0.23402
36	0.15654	0.48923	-0.05420	-0.10608	0.30206	0.00869	0.03037
37	0.24239	0.19647	0.07886	-0.06020	0.14351	-0.01214	0.29148
38	0.08278	0.34292	0.09161	-0.36513	0.28305	-0.05409	-0.10102
39	0.01767	0.87804	-0.01288	0.03490	0.01106	0.01860	0.03974
40	-0.08031	0.85101	0.00846	0.02022	0.02992	0.06365	0.10131

TABLE VI

LANGUAGE ASSESSMENT BATTERY
ENGLISH - LEVEL 1 - TOTAL SCORE
PUBLISHED NORM DATA

	Grade		
	K	1	2
Mean	24.40	32.59	37.60
S.D.	4.24	5.11	3.76
KR-20	0.82	0.86	0.89
Standard Error Measure	1.80	1.92	1.24

TABLE VII

95% CONFIDENCE INTERVALS

Above and Below 23%ile

Grade	To Be Above 23%ile		To Be Below 23%ile	
	R.S.	%ile	R.S.	%ile
K	25	66	16	2
1	32	49	24	5
2	39	57	33	13

Figure 1
LANGUAGE ASSESSMENT BATTERY
PUBLISHED NORM TABLES PLOTTED
ENGLISH - LEVEL 1

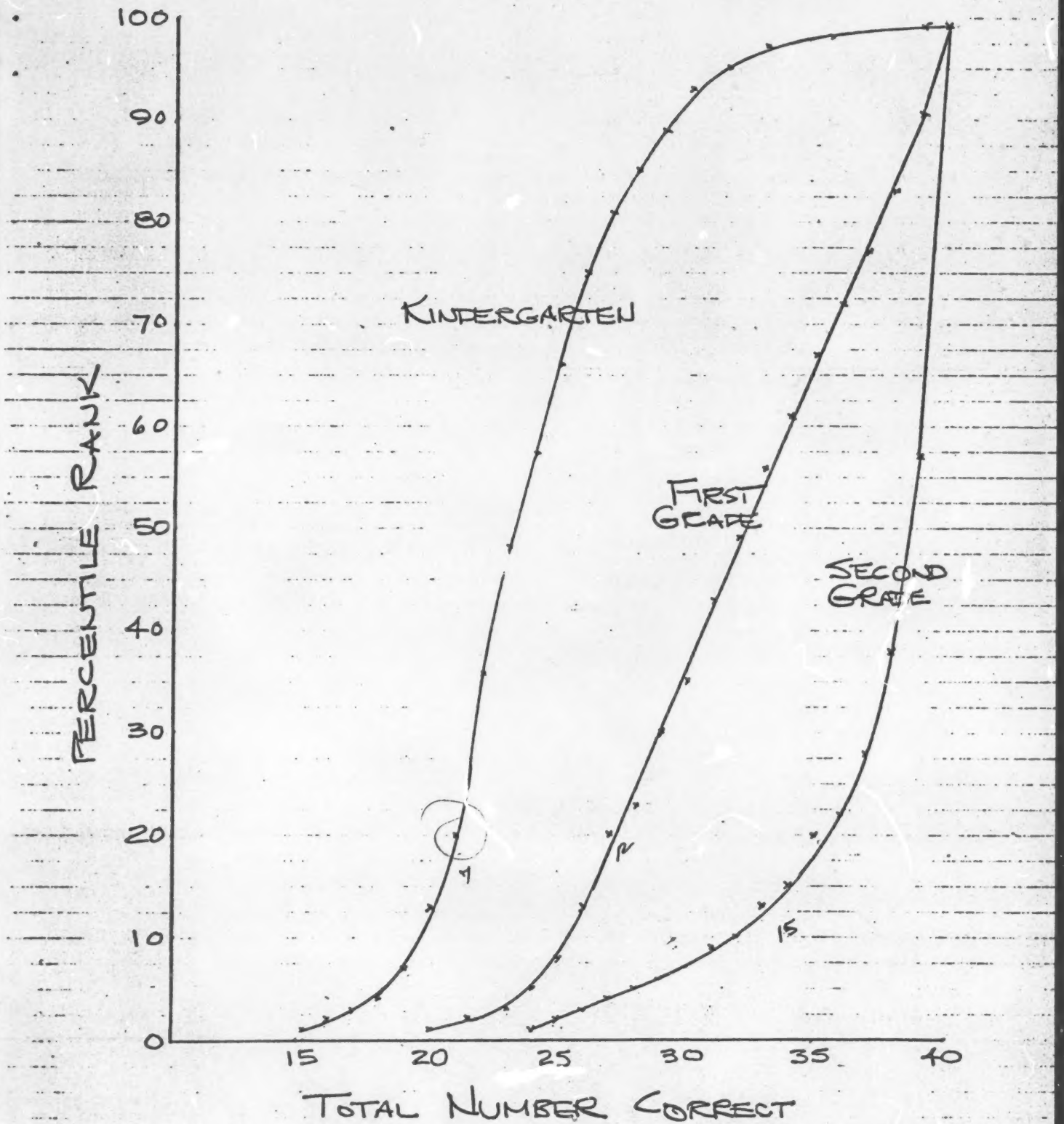
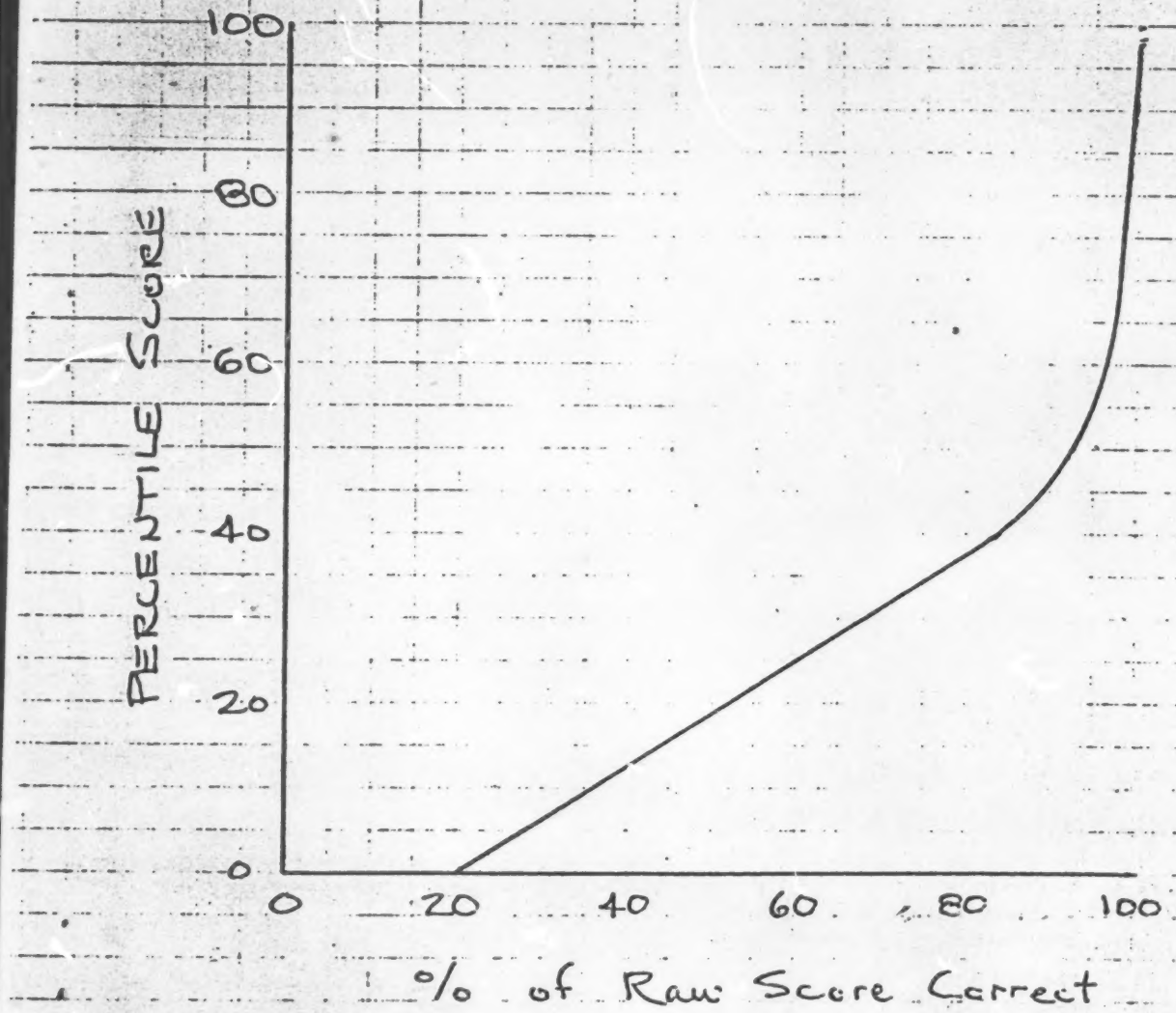


FIGURE 2

IDEAL NORMS FOR LANGUAGE ASSESSMENT BATTERY



Bibliography

Hernandez-Chavez, E.; Burt, M. & Dulay, H. Language dom.nance and proficiency testing: some general considerations. NABE Journal, Fall 1978, Vol III No. 8, 41-54.

Kerlinger, F.N.; & Pedhazur, F.N. Multiple Regression in Behavioral Research. New York: Holt, Rinehart, and Winston, Inc. 1973.

Kim, J. Factor Analysis. In Statistical Package for the Social Sciences (2nd Ed.) Nie, Hull, Jenkins, Steinbrenner, Bent. New York: Mcgraw-Hill Book Co., 1975.

MacNamara, J. The bilingual's linguistic performance- A psychological overview. Journal of Social Issues Vol. 23, No. 2, 58-77.

Nunnally, J.C. Psychometric Theory. New York: McGraw-Hill Book Co. 1967.